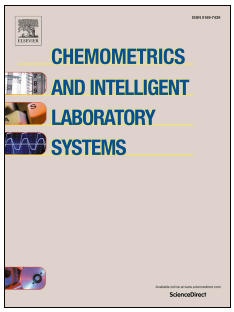
Visualising interactions in bi- and triadditive models for three-way tables

Accepted Manuscript



Visualising interactions in bi- and triadditive models for three-way tables

Casper Albers, John Gower

PII: S0169-7439(17)30323-4

DOI: 10.1016/j.chemolab.2017.05.014

Reference: CHEMOM 3445

To appear in: *Chemometrics and Intelligent Laboratory Systems*

Received Date: 3 September 2015

Revised Date: 12 May 2017

Accepted Date: 15 May 2017

Please cite this article as: C. Albers, J. Gower, Visualising interactions in bi- and triadditive models for three-way tables, *Chemometrics and Intelligent Laboratory Systems* (2017), doi: 10.1016/j.chemolab.2017.05.014.

Visualising interactions in bi- and triadditive models for three-way tables

Casper Albers* John Gower†

May 16, 2017

Abstract

This paper concerns the visualisation of interaction in three-way arrays. It extends some standard ways of visualising biadditive modelling for two-way data to the case of three-way data. Three-way interaction is modelled by the Parafac method as applied to interaction arrays that have main effects and biadditive terms removed. These interactions are visualised in three and two dimensions. We introduce some ideas to reduce visual overload that can occur when the data array has many entries. Details are given on the interpretation of a novel way of representing rank-three interactions accurately in two dimensions. The discussion has implications regarding interpreting the concept of interaction in three-way arrays.

Keywords. Interpretation of interaction, Modelling of interaction, Visualisation of interaction, Biadditive Models, Individual Scaling.

*Department of Psychometrics & Statistics, University of Groningen, Groningen, The Netherlands; corresponding author, c.j.albers@rug.nl.

†Department of Mathematics & Statistics, The Open University, Milton Keynes, United Kingdom.

1 Setting the scene

> "...It is important that the final model or models should make sense physically: at a minimum, this usually means that interactions should not be included without main effects nor higher-degree polynomial terms without their lower-degree relatives. Furthermore, if the model is to be used as a summary of the findings of one out of several studies bearing on the same phenomenon, main effects would usually be included whether significant or not. Strict adherence to this policy makes it easier to compare the results of various studies and helps to avoid the apparent conflicts that occur when different fitted models with different sets of terms are used in each study."
>
> McCullagh and Nelder (1989, p.89)

In this paper, we are concerned with three-way tables $\underline{\mathbf{X}}$ with elements x_{ijk} ($i = 1, \ldots, I$; $j = 1, \ldots, J$; $k = 1, \ldots, K$). Thus, the factors used to classify the three ways have equal status (sometimes called modes) while the body of the table contains values of a quantitative variable that may be regarded as a dependent variable - as classically typified by a three-way table arising from agricultural experiments with fertilizer treatments as factors and crop yield as the response. The factors are treated as categorical variables but if they happen to have numerical values, this may be taken into account when interpreting interactions. The primary emphasis is on the visualisation of interaction with a supplementary interest in estimation and interpretation seen in the light of the quotation from McCullagh and Nelder (1989). To dispel any suggestion to the contrary, we emphasize that the quotation is not an expression of a mathematical fact but more an observation on how data can usually be expected to behave. In the psychometric literature, a three-way table is sometimes referred to as one-mode three-way data (Carroll and Arabie, 1980; Coombs, 1964; Kiers, 2000) or, shorter, as (data) array, whereas in chemometrics the terminology tensor for $\underline{\mathbf{X}}$ is more common.

Three-way tables are usually analysed by linear models containing additive terms representing main effects, two-factor interactions, and three-factor interactions. The number of factors can be readily extended to any number of "ways". The form of such models readily respects the McCullagh and Nelder (1989) quotation. Note that with a dependent interval variable there is a fundamental need for at least one additive parameter to represent translation (e.g. Celsius to Fahrenheit).

For reference, and to establish notation, we list the basic results for additive models. The model is

$$x_{ijk} = m + \{a_i + b_j + c_k\} + \{d_{jk} + e_{ik} + f_{ij}\} + g_{ijk} \tag{1}$$

where the terms with a single suffix represent main effects, those with double suffices two factor interactions and g_{ijk} represents contributions from three factor interactions. Some components of the interactions may be regarded as "error". The estimating equations are subsumed in the identity:

$$\begin{aligned}\hat{x}_{ijk} &= x_{...} + \{(x_{i..} - x_{...}) + (x_{.j.} - x_{...}) + (x_{..k} - x_{...})\} \\ &\quad +\{(x_{.jk} - x_{.j.} - x_{..k} + x_{...}) + (x_{i.k} - x_{i..} - x_{..k} + x_{...}) \\ &\quad +(x_{ij.} - x_{i..} - x_{.j.} + x_{...})\} \\ &\quad +(x_{ijk} - x_{.jk} + x_{i.k} + x_{ij.} + x_{i..} + x_{.j.} + x_{..k} - x_{...})\end{aligned} \tag{2}$$

where the expressions in braces in (2) estimate the corresponding parameters in (1). Note that we adopt the convention that a "hat" on the left-hand-side implies that the terms on the right-hand-side are parameter estimates, else they are the parameters themselves. The terms in (2) contribute to an orthogonal analysis of variance:

$$\sum_{i,j,k}^{I,J,K} (\hat{x}_{ijk} - x_{...})^2 = JK||\mathbf{a}||^2 + IK||\mathbf{b}||^2 + IJ||\mathbf{c}||^2 + I||\mathbf{D}||^2 + J||\mathbf{E}||^2 + K||\mathbf{F}||^2 + ||\mathbf{G}||^2 \tag{3}$$

where $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are vectors of the main effects, $\mathbf{D}, \mathbf{E}, \mathbf{F}$ are matrices of the two-factor interactions and $||\mathbf{G}||^2$ represents the sum-of-squares of the elements of the three-factor interaction.

When interactions have been estimated, there remains the problem of their interpretation. The terms in (2) represent overall contributions to each main effect and interaction. To help interpret overall representations of interaction, several simple approximations have been proposed. One possibility is to focus on the larger (positive or negative) terms. Another is to fit linear and quadratic polynomials to get, for example, linear × linear × quadratric estimates. Even the simpler of these can be difficult to interpret and, strictly speaking, such expressions are valid only when the classifying factors are numerical (like levels of fertilizer applications).

Another possibility is to fit product terms like $a_i b_j$. Products of two factors have bilinear regression interpretations and a nice geometrical representation that underpins useful visualisations of two-factor interaction. This possibility of biadditive modelling is discussed in Section 2. A biadditive model gives the best rank-r least-squares approximation to a two-way table/matrix but this optimal mathematical property should not necessarily be taken as an expression of an appeal to underlying substantive multiplicative effects.

In a parallel literature, models for analysing three-way data (summarised in Kroonenberg, 2008; Smilde et al., 2004) often include triple product terms like $a_i b_j c_k$. Included are three-mode principal component analysis (Tucker, 1966) and methods as the Candecomp (Carroll and Chang, 1970) and Parafac models (Harshman, 1970) (both models are equivalent and commonly denoted as the CP-model). A desirable computational requirement for fitting three-way multiplicative models is a universal algorithm for fitting a general canonical decomposition for three-way arrays. Such models are discussed in Section 3. It is clear that triple product terms may be potentially useful in many contexts and considered as a natural extension for representing triadditive interactions in a similar way that biadditive models may represent two-factor interactions.

In many psychometric and chemometric methods, the triple product term dominates the model, even to the extent of excluding lower order terms, thus not respecting the maxim of McCullagh and Nelder (1989) cited at the start of this paper. This is because in psychometrics the methods are intended as generalisations of Principal Component Analysis and related methods that do not admit a dependent variable; such methods are beyond the scope of this paper. Nevertheless, triadditive terms may be used to approximate three-way interactions. In the following we exploit the fact that the Candecomp-Parafac algorithm can be useful for fitting three-way multiplicative interactions in three-way models. We explore the consequences for the McCullagh and Nelder dictum if this route is taken. Visualisation is important in the interpretation of biadditive interactions and we provide suggestions for its improvement: Appendix A discusses how to calibrate axes Appendix B provides details on optimising a parallel axis display of the interactions and Section 4 demonstrates these methods. Furthermore, we show how triadditive terms may be visualised and interpreted.

In the above, we have regarded the overall main effects and interaction terms in (2) as the definitive expressions of interaction. These may then be approximated as we have described, by linear, biadditive or triadditive estimates, perhaps including other parts of the interactions in an error term. For linear and biadditive estimates the procedure of estimating the biadditive part of each interaction, conditionally on the usual least-squares estimates of the linear part, usually turns out to be equivalent to unconditional estimation. However, this is not true for some of the biadditive models we discuss below and for triadditive models it is never true.

Sections 2 and 3 briefly summarise some of the current insights in biadditive and triadditive models and discuss various ways of modelling and interpreting interactions using these models. These sections are not meant provide an exhaustive and complete overview of all knowledge on biadditive and triadditive models, as good sources for that already exist (Smilde et al., 2004; Kroonenberg, 2008). Subsequently, biadditive and

triadditive visualisations are constructed for an example from agricultural (Section 4) research. Although these visualisations are based on the Candecomp-model Carroll and Chang (1970), the visualisations can also be based on other techniques for analysing three-way arrays. Section 5 concludes the paper.

2 Biadditive models

In this section we summarise well-known results for biadditive models. This establishes notation that is needed for similar developments with triadditive models discussed in Section 3.

2.1 Biadditive models for two-way tables

For an $I \times J$ table $\underline{\mathbf{X}}$ with elements x_{ij} the general biadditive model is:

$$x_{ij} = m + a_i + b_j + \sum_{r=1}^{R} c_{ir}\tilde{c}_{jr} + \varepsilon_{ij} \quad (i = 1, \ldots, I, j = 1, \ldots, J) \tag{4}$$

where a_i and b_j represent row and column main effects, and c_{ir} and $\tilde{c}_{jr}$ $(r = 1, \ldots, R)$ model the multiplicative interaction. The error terms ε_{ij} are assumed to be independently distributed with equal variances. Many classical models, such as Tukey's model for one degree of freedom for non-additivity (Tukey, 1949), can be considered as special cases of a biadditive model. Alternative names under which (4) has appeared, are FANOVA (FActor ANalysis Of VAriance) (Gollob, 1968) and AMMI (Additive Main effects and Multiplicative Interactions) (Gauch, 1992). Also the GEMANOVA (Generalised multiplicative ANOVA) model (cf. Bro and Jakobsen, 2002) is related. We prefer the neutral biadditive model terminology which is in line with general statistical usage (Denis and Gower, 1994). These authors were interested in biadditivity because they thought that substantive genetic effects were better modelled in multiplicative rather than additive terms.

In general, model (4) is not fully identified. The simplest identification constraints for the general model are

$$\left.\begin{array}{l} \mathbf{1}'\mathbf{a} = \mathbf{1}'\mathbf{b} = \mathbf{1}'\mathbf{c}_r = \mathbf{1}\tilde{\mathbf{c}}_r = 0 \\ \tilde{\mathbf{c}}_r'\tilde{\mathbf{c}}_r = \mathbf{c}_r'\mathbf{c}_r = \sigma_r, \text{ say, and } \tilde{\mathbf{c}}_r'\tilde{\mathbf{c}}_s = \mathbf{c}_r'\mathbf{c}_s = 0 \ (r \neq s) \end{array}\right\} (r, s = 1, \ldots, R) \tag{5}$$

ensuring that the matrix $\sum_{r=1}^{R} c_{ir}\tilde{c}_{jr}$ of interaction parameters of rank R is uniquely parameterised in the form of its singular value decomposition with singular values

$\sigma_1, \ldots, \sigma_R$.

The analysis of variance corresponding to a two-way version of (4) is:

$$\sum_{i,j}^{I,J} (\hat{x}_{ij} - x_{..})^2 = J||\mathbf{a}||^2 + I||\mathbf{b}||^2 + \sum_{r=1}^{R} \sigma_r^2 + \sum_{r=R+1}^{\rho} \sigma_r^2 \tag{6}$$

where $\rho = \text{rank}(\underline{\mathbf{X}})$.

2.2 Biadditive models for three-way tables

Biadditive terms may be used to model interaction in three-way tables (cf. Gower, 1977). For an $I \times J \times K$ table $\underline{\mathbf{X}}$ with elements x_{ijk} we may consider the following biadditive model:

$$x_{ijk} = m + a_i + b_j + c_k + \sum_{p=1}^{P} d_{jp}\tilde{d}_{kp} + \sum_{q=1}^{Q} e_{iq}\tilde{e}_{kq} + \sum_{r=1}^{R} f_{ir}\tilde{f}_{jr} + \varepsilon_{ijk} \tag{7}$$

for $i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K$, where the ε_{ijk} are the elements of the three-way error array $\underline{\mathbf{E}}$.

Similar identification constraints to those already discussed for model (4) may be applied for the biadditive model (7) for three-way tables:

$$\mathbf{1}'\mathbf{a} = \mathbf{1}'\mathbf{b} = \mathbf{1}'\mathbf{c} = \mathbf{1}'\mathbf{d}_p = \mathbf{1}'\tilde{\mathbf{d}}_p = \mathbf{1}'\mathbf{e}_q = \mathbf{1}'\tilde{\mathbf{e}}_q = \mathbf{1}'\mathbf{f}_r = \mathbf{1}'\tilde{\mathbf{f}}_r = 0$$

for $p = 1, \ldots, P$; $q = 1, \ldots, Q$; $r = 1, \ldots, R$, together with the SVDs of the three biadditive interaction matrices as they occur in (2). The resulting analysis of variance is:

$$\sum_{i,j,k=1}^{I,J,K} (\hat{x}_{ijk} - x_{...})^2 = JK||\mathbf{a}||^2 + IK||\mathbf{b}||^2 + IJ||\mathbf{c}||^2 + I\sum_{s=1}^{P}\sigma_{ps}^2 + J\sum_{s=1}^{Q}\sigma_{qs}^2 + K\sum_{s=1}^{R}\sigma_{rs}^2 + \sigma^2 \tag{8}$$

where the singular values σ_{ps} $(s = 1, \ldots, P)$, σ_{qs} $(s = 1, \ldots, Q)$ and σ_{rs} $(s = 1, \ldots, R)$ refer to the respective residual tables $\underline{\mathbf{Z}}_i$, $\underline{\mathbf{Z}}_j$ and $\underline{\mathbf{Z}}_k$ defined as in (2), and σ^2 is the residual sum-of-squares obtained from all the singular values not included in the summations. The solution for the multiplicative constants is then obtained from the SVD of the two-way tables of residuals $\underline{\mathbf{Z}}_i$, $\underline{\mathbf{Z}}_j$ and $\underline{\mathbf{Z}}_k$. This is a simple generalisation that may be readily extended to tables of any number of "ways".

The choice of ranks P, Q and R can be made by ad hoc arguments, such as that rank 2 approximations can be visualised and communicated in an understandable way.

Another option lies in more formal arguments such as obtaining corresponding degrees of freedom, for instance for the A× B interaction, through the rule of thumb that (i) degrees of freedom for $P = 1, 2, \ldots, \min(I-1, J-1)$ should add up to that of the A×B interaction in the two-way ANOVA table, (ii) the df for dimension i should be two less than that for dimension $i-1$. According to (Gower et al., 2011, Section 6.3), this rule was first given by Rao (1952). A formal test of significance for P, Q or $R = 1$ has been given by Corsten and Eijnsbergen (1972). Other approaches include cross-validation and using multiway extentions of the Kaiser criterion or scree plot (Kroonenberg and van der Voort, 1987), such as the DifFit procedure for Tucker3 models (Timmerman and Kiers, 2000). See Smilde et al. (2004, Section 7.4) and Kroonenberg (2008, Section 8.5) for an overview of component-selection methods.

2.3 Visualisation for biadditive models

It is useful, especially when $R = 2$, to plot the rows of $\mathbf{c}_r$ $(r = 1, \ldots, R)$ to give I row-points and the rows of $\tilde{\mathbf{c}}_r$ $(r = 1, \ldots, R)$ to give J column-points. In this biplot, the inner-product determined by a pair of points, one from each set, gives a visualisation of the corresponding interaction. This is a well-known form of biplot (see e.g. Gower et al., 2011). Another possibility is to present the rows as axes and the columns as points (or vice versa). The axes may be calibrated, making it trivial to find values of inner products.

Furthermore, axes may include markers for the row or column main effects. As we show in Appendix A, calibrated axes may be provided *simultaneously* for rows and columns *and* both sets of main effects may be included. In addition, the values of $\alpha + \beta = 1$ (as defined in Appendix A) are at choice and λ-scaling is available (see Gower et al., 2011). In this way, a variety of equivalent representations, which may be regarded as items drawn from a toolbox, is available for presentational purposes. One may choose among the possibilities to represent only the more important interactions. Some examples are included in Section 4 of this paper.

The biplot representation of two-factor interactions is an attractive aid to interpretation. Also the biadditive model of three-way data can be visualised, now by three biplots, one for each biadditive term in (7).

3 Triadditive models

3.1 Triadditive models for three-way data

For an $I \times J \times K$ table $\underline{\mathbf{X}}$ with elements x_{ijk}, consider the following triadditive model:

$$x_{ijk} = m + a_i + b_j + c_k + \sum_{p=1}^{P} d_{jp}\tilde{d}_{kp} + \sum_{q=1}^{Q} e_{iq}\tilde{e}_{kq} + \sum_{r=1}^{R} f_{ir}\tilde{f}_{jr} + \sum_{s=1}^{S} g_{is}\tilde{g}_{js}\tilde{\tilde{g}}_{ks} + \varepsilon_{ijk} \quad (9)$$

This model is an extension of (7) where the error array $\underline{\mathbf{E}}$ is partitioned into a rank-S triadditive part $\mathbf{G}$ and a new error array $\underline{\mathbf{E}}$ with, generally, a smaller sum of squared elements than that of (7). For identification, the usual zero-sum identification constraints may be applied to all the parameters but when applied to the triadditive parameters g_{is}, $\tilde{g}_{js}$, $\tilde{\tilde{g}}_{ks}$ it has unexpected implications. This is because adding constants α, β, γ replaces the triadditive term by $(g_{is}+\alpha)(\tilde{g}_{js}+\beta)(\tilde{\tilde{g}}_{ks}+\gamma)$ which, on expansion, induces additional additive and biadditive terms. The additive terms may be absorbed into zero-sum main effects without affecting the form of the model. This is not so for the biadditive terms, where unabsorbable parts of the triadditive interaction contribute to the biadditive parameters, thus increasing their rank. Thus, this reparameterisation changes the form of the model. One consequence is that the least-squares estimates of the triadditive interaction parameters are not the same as the estimates conditional on the estimated main effects and biadditive interactions. Another, is that the usual orthogonal analysis of variance is not available. This position may be accepted and algorithms developed to fit the model but a more simple option is to fit the triadditive part conditional on the main effects and the saturated biadditive component of the model. That is, we fit the triadditive part of the model to the biadditive residual table:

$$\hat{z}_{ijk} = x_{ijk} - x_{\cdot jk} - x_{i\cdot k} - x_{ij\cdot} + x_{i\cdot\cdot} + x_{\cdot j\cdot} + x_{\cdot\cdot k} - x_{\cdots}. \quad (10)$$

Triadditive interactions in (9) may be modelled in two ways. If z_{ijk} represents a typical term of the interaction we may fix one factor, say i, and consider the I two-way tables $\{z_{1jk}\}, \{z_{2jk}\}, \ldots, \{z_{Ijk}\}$. Each of these tables may be fitted by a biadditive model and the results compared. This approach is consistent with the classical notion of interaction as a difference in response to a factor, or set of factors (here j and k), at different levels of another factor (here i). Of course, we may interchange the roles of i, j and k. The other approach is to fit a truly triadic model with the Candecomp-Parafac

algorithm (Carroll and Chang, 1970; Harshman, 1970), minimising:

$$\sum_{i,j,k=1}^{I,J,K} \sum_{r=1}^{R} \left(z_{ijk} - u_{ir} v_{jr} w_{kr}\right)^2 . \tag{11}$$

We choose for this approach as it is a truly triadic approach. The approximation (11) may be viewed as the triadditive counterpart of the Eckart-Young theorem but lacking a nice known canonical decomposition. (See also Schmidt (1907), which is said to be the first example of the SVD least-squares property, albeit in a very different field from data analysis.) This approach is close to the classical approximation of interactions by orthogonal polynomials in linear models. Here we fit a biadditive approximation to the two-way interactions and a triadditive approximation to the three-way interaction terms of (1) and (2). The residuals from the triadic term contribute to the term (11), while the biadditive part contributes components what we denote by σ^2 in (8). In a good fit, these two components should be comparable giving some indication of stability and, when available, they may be compared with independent estimates of replication-error. From the statistical point of view we need some concept akin to that of degrees of freedom in linear models. What is known about this is summarised by Kroonenberg (2008, Section 8.4). Related to this is the concept of rank for three-way arrays (cf. ten Berge (2011) and Smilde et al. (2004, Section 2.6)). Triadditive rank is defined as the smallest value of R that gives an exact triadditive fit. The interaction array $\underline{\mathbf{Z}}$, with its zero marginals, generally has lower rank than the data array $\underline{\mathbf{X}}$ (Albers et al., 2017). Since our focus lies on the visualisation of interactions, here we will not formally study rank properties of $\underline{\mathbf{Z}}$.

3.2 Visualisation for three-way data

As with the biadditive model, when a rank R triadic model (11) has been fitted, there is interest in expressing the interaction in graphical form. In the rank one case ($R = 1$), the points for u_{i1} ($i = 1, \ldots, I$); v_{j1} ($j = 1, \ldots, J$); w_{k1} ($k = 1, \ldots, K$) may be placed on separate orthogonal coordinate axes, which we shall label u, v and w. Then, $u_{i1} v_{j1} w_{k1}$ is simply proportional to the volume of the tetrahedron with these three points on orthogonal *axes* and the origin as vertices (Figure 1, left).

When $R = 2$, the visualisation remains basically Euclidean in three dimensions and it may be interpreted in terms of tetrahedronal volume where the vertices of the tetrahedra are confined to the origin and three orthogonal *planes* (Figure 1, right). The justification

of this approach follows from the trilinear identity:

$$\det \begin{pmatrix} 0 & u_{i1} & u_{i2} \\ v_{j2} & 0 & v_{j1} \\ w_{k1} & w_{k2} & 0 \end{pmatrix} = u_{i1}v_{j1}w_{k1} + u_{i2}v_{j2}w_{k2} \tag{12}$$

(see also equation (4) in Albers and Gower (2014)). The rows of the determinant on the left hand side may be interpreted as giving the coordinates of three points, one in each of three orthogonal dimensions, while the right hand side gives a term in the rank two triadditive model. Albers and Gower (2014) give further details and show that, without loss of information, this representation may be shown in two dimensions to give a visualisation which resembles a biplot, with one set of K coplanar points and two sets of calibrated axes representing the remaining IJ factors. Thus, it is a 'triplot' rather than a biplot (see e.g. Gower et al., 2011). Whilst Albers and Gower (2014) explain the technical construction of these triplots, instruction on how to interpret these triplots, especially in the case of interaction arrays, is lacking. We provide such explanation Section 4. That rank-two trilinear interactions may be shown in two dimensions, gives them similar status to interactions for bilinear models and makes direct three-dimensional tetrahedral visualisations unnecessary. We believe that this is a major step forward.

* FIGURE 1 ABOUT HERE *

Because volume is invariant to orthogonal transformations, one may deduce from the above three-dimensional representation that the parameters of rank 2 triadditive models are determined only up to arbitrary orthogonal rotations in three dimensions. This degree of arbitrariness is similar to that found in biadditive models where inner-products or, equivalently, areas (Gower et al., 2010) rather than volume are the invariants. Orthogonal transformation is not the only invariant for rank 2 triadditive models; for example, provided $\alpha\beta\gamma = 1$, we could also scale the three axes by α, β, γ, respectively, without affecting volume. Our experience is that visualisation that yields easiest interpretation is achieved when α, β and γ are chosen such that $\sum_{i,r} u_{i,r}^2 \approx \sum_{j,r} v_{j,r}^2 \approx \sum_{k,r} w_{k,r}^2$. With this degree of arbitrariness, we see little point in paying much attention to the estimated values of the parameters u, v, w but rather to focus on the invariants, such as volume and the actual fitted values $\hat{x}_{ijk}$ and $\hat{z}_{ijk}$.

Higher rank solutions to biadditive models can be shown as three-dimensional images or by exhibiting several planar cross-sections of the higher-dimensional space. Neither of these is satisfactory and it is the two-dimensional approximations that are by far the most important. Nevertheless, it is interesting to see what progress can be made with

representing triadditive terms for $R = 3$. We could show this as three volumes, each of unit rank $(u_{i1}v_{j1}w_{k1}) + (u_{i2}v_{j2}w_{k2}) + (u_{i3}v_{j3}w_{k3})$, or of two volumes, one of unit rank and the other of rank two $(u_{i1}v_{j1}w_{k1}) + (u_{i2}v_{j2}w_{k2} + u_{i3}v_{j3}w_{k3})$. A more symmetric representation arises from noting that

$$\begin{aligned} & 2(u_{i1}v_{j1}w_{k1} + u_{i2}v_{j2}w_{k2} + u_{i3}v_{j3}w_{k3}) \qquad (13) \\ = \ & \det\begin{pmatrix} 0 & u_{i1} & u_{i2} \\ v_{j2} & 0 & v_{j1} \\ w_{k1} & w_{k2} & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & u_{i1} & u_{i3} \\ v_{j3} & 0 & v_{j1} \\ w_{k1} & w_{k3} & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & u_{i2} & u_{i3} \\ v_{j3} & 0 & v_{j2} \\ w_{k2} & w_{k3} & 0 \end{pmatrix}. \end{aligned}$$

After equation (12), we explained how this determinant is equal to the volume of a single tetrahedron. Using analogous arguments, equation (13) equals three times the sum of the volumes of the tetrahedra designated by the three separate determinants. We have seen that when $R = 1$, the three axes share a common origin and when $R = 2$ the three planes share an orthogonal set of axes. When $R = 3$ we retain the orthogonal axes u, v, w but, as is shown by (13), it is the projections of the points $(u_{i1}u_{i2}u_{13})$, $(v_{j1}v_{j2}v_{j3})$, (w_{k1}, w_{k2}, w_{k3}) onto the (vw), (wu), (uv) planes that determine the vertices of the operative tetrahedra. The display of Figure 2 shows that this visualisation is on the boundary of what is relevant for practical purposes.

Interestingly, when $R = 4$ we may write $(u_{i1}v_{j1}w_{k1} + u_{i2}v_{j2}w_{k2}) + (u_{i3}v_{j3}w_{k3} + u_{i4}v_{j4}w_{k4})$ the sum of two rank 2 terms each representable by a single tetrahedron. However, adding even two volumes is not acceptable. We conclude that rank two representations of triadditive models are at the limits of useful graphical representation; higher ranks are possible but are impracticable.

* FIGURE 2 ABOUT HERE *

4 Application: response of wheat varieties to the application of nitrogen fertiliser at different sites

Blackman et al. (1978) studied the effect of the application of nitrogen fertiliser to several varieties of winter wheat of contrasting height grown at different trial sites. The data consists of a fully crossed design with the following three factors:

A Rate of nitrogen application ($I = 2$ levels, low and high)

B Trial sites ($J = 7$ locations in the United Kingdom)

C Variety ($K = 12$ different varieties).

The names of the factor levels for factors B and C are given in Table 1. A fourth factor, indicating whether the variety is either 'conventional' (varieties Cappelle, Ranger, Huntsman, Templar, and Kinsman) or 'semi-dwarf' (varieties Fundin, Durin, Hobbit, Sportsman, TJB295/95, TJB325/464, and Hustler), is excluded from our analysis as it's obviously not a crossed factor. The dependent variable is grain yield, measured in grams per square meter. One trial site (Edinburgh) is located in Scotland, the six others are all located in Cambridgeshire and Oxfordshire, England. In this section we are mainly concerned with visual presentation of interactions rather than with substantive analysis.

4.1 Biadditive visualisation

First, we fit the biadditive model as outlined in Section 2.2. Table 2 shows that factor B, Trial Site, is the most important main factor and the interaction between A, rate of nitrogen application, and B is the most important two-way interaction. The main effects constitute 84% of total variation in grain yield, the two-way interactions 14% and the three-way interaction 2%.

Table 2 also provides the sums-of-squares of the low-rank approximations to the two-way interaction between B and C, according to Equation (7) with approximations to degrees of freedom as suggested by Rao (1952) (see Section 2.2). Since Factor A has two levels, $df_A = 1$. Hence, this low-rank approximation does not apply to the AB and AC interactions: the full-rank approximation is already of the lowest rank possible. Were $df_A > 1$, the treatment of the low-rank approximations to interactions AB and AC would have been analogous to that of BC. Corresponding to BC, most information, 79%, is captured in the first two dimensions.

For this data, two-dimensional biplots of interactions with factor A are not relevant: A has only two levels, thus the interactions are one-dimensional. Figure 3 gives a series of equivalent biplots for interaction BC. In all cases, interpretation is through evaluating inner-products, either directly or indirectly. Figure 3a visualises the interaction BC in the conventional way. Often, the points are connected to the origin and perhaps endowed with arrows. The interactions of the varieties at the trial site in Edinburgh clearly deviate from those at the six English sites. A closer examination confirms that the McCullagh and Nelder dictum, cited at the beginning of this paper, holds. Interestingly, no clear distinction in interaction can be found between the regular and the semi-dwarf varieties.

Figure 3a is useful for assessing global patterns in the data but no numerical values can be read off. For this, calibrated axes are needed. The technicalities behind the construction of such axes simultaneously for sites and varieties is explained in Appendix A. The biplots in the other panels make use of such calibrated axes. They give the same

information as Figure 3a, but in 3b and 3c, while varieties continue to be represented by points, trial sites are represented by calibrated axes. The Figures show exclusion (3b) vs. inclusion (3c) of main effects but otherwise are identical; thus Figure 3b displays the biadditive interactions *after* the main effects have been partialed out, whereas these are still included in Figure 3c. The only difference between panels (b) and (c) is the calibration of the axes: where in panel (b) all axes have value 0 at the origin, this is not the case in panel (c). Figure (3d) shows calibrated axes for both varieties and sites. Note that a variety projected onto a site-axis gives the same calibration as the same site projected onto the corresponding variety axis. For example, consider variety Sportsman and site Edinburgh (as shown in Figure 3(d)): The projection of Sportsman onto Edinburgh is -30.33 g/sqm, which is equivlent to the projection of Edinburgh onto Sportsman. The same holds for all other pairs of sites and varieties.

* TABLES 1 & 2, FIGURES 3 & 4 ABOUT HERE *

Thus, with Figure 3(a) inner products are not needed to rank varieties *within* a site or to rank sites growing the same variety but it is difficult to make numerical comparisons *between* sites and varieties. This problem is reduced by using the calibrations in Figure 3(b) and Figure 3(c) but the calibration markers tend to lead to problems of visual overload.

Figure 4 is a compromise which preserves most of the useful information and is easy to use. Essentially, it consists of taking the axes of one set of calibrations (say, the seven sites) and laying them horizontally on successive lines with a common origin in a so-called parallel coordinate plot (cf. Inselberg, 2009). The different interval of calibration on each axis will be clear and can be removed by normalising each line to have an equal interval of calibration. Then, the calibration markers on the successive lines can be removed and replaced by a single calibrated axis applicable to all sites, as shown in Figure 4. Parallel coordinate plots date back to (at least) the 17th century (d'Ocagne, 1885) and gained popularity through the work of Inselberg in the past four decades (Inselberg, 2009). The usage of parallel coordinate plots in the context of three-way analysis is not new (cf. Kroonenberg, 2008, p. 400), but this paper is, to our knowledge, the first that employs parallel coordinate plots to visualise three-way interactions.

In this example, there is no logical ordering for the sites. Rather than the alphabetical ordering in Figure 4, any other of the $J! = 5040$ orderings can be used. Although all variations provide exactly the same information, some allow for easier interpretation because there is less 'clutter', such as fewer line-crossings. When J is not too large, one can resort to manual reordering but for larger values of J, an automated procedure is preferable. We propose such a procedure, based on correspondence analysis (cf.

Greenacre, 2007). Technicalities of this procedure are provided in Appendix B and Figure 5 shows the optimal ordering. This figure provides *exactly* the same information as Figure 4 but is easier to interpret.

Now, the performance of every variety at each site may be readily compared directly. The main effects may be included if desired but we have not done this with these data because of the disproportionate main effect of Edinburgh (a value +232 grams per square meter; whereas the other six sites have main effects between −113 and +46 grams per square meter). Of course, an equivalent procedure can be used for the varieties rather than for the sites.

4.2 Triadditive visualisation

Having eliminated all main and multiplicative effects according to (10), Candecomp-Parafac approximations of different rank were fitted to $\hat{\underline{\mathbf{Z}}}$. Table 3 displays the breakdown of the ABC-interaction SS of 49812 into approximations of rank 1 to 6. Rank 2 and 3 approximation explain 63% and 78% of the variation in grain yield, respectively. Thus, visualisations on the basis of these approximations will yield useful insight into the structure of $\hat{\underline{\mathbf{Z}}}$.

Figure 1 visualises the rank 1 and 2 approximations to the three-way interaction term. The highlighted interaction in each figure is that between a low rate of nitrogen application, trial site Edinburgh variety Kinsmen. The data have been scaled by α, β, and γ in such a way that

$$\sum_{i=1}^{I}\sum_{r=1}^{R} u_{ir}^2 = \sum_{j=1}^{J}\sum_{r=1}^{R} v_{jr}^2 = \sum_{k=1}^{K}\sum_{r=1}^{R} w_{kr}^2$$

because this provided a satisfactory visual setting for interpretation (the dispersion in the three dimensions is made the same; without affecting the volume of the tetrahedra). Figure 1(left) visualises the rank $R = 1$ approximation and shows how, by looking at tetrahedra, one can quickly get an impression of a specific triadditive interaction. Figure 1(right) displays the visualisation for $R = 2$, via three biplots for the three factors. Each biplot may be visualised in one of the three orthogonal planes (uv, uw and vw) through the origin. The interaction of interest remains proportional to the volume of a single tetrahedron.

A three-dimensional rank $R = 3$ visualisation is crossing the line of useful application (as outlined in Section 3.2). It is much more simple to look at a two-dimensional visualisation through a so-called triplot. Here, we use the term triplot in the same

way as in Albers and Gower (2014). According to Williams and Gardner-Lubbe (2016), the use of the term 'triplot' in this context dates back to Araújo (2009). Meulman et al. (2004, p. 50) also use this term, in a slightly different context related to biplots. Furthermore, the term triplot is also used in for triangular diagrams, which is a unrelated field of work. As the contexts are fully different, this should not cause confusion.) In this display, each IJ combination of levels is represented by a calibrated axis while each level of K is represented by a point (for the Blackman data we have $I = 2$, $J = 7$ and $K = 12$). Thus, an axis combining a Site (e.g. Edinburgh) with the Higher Level of Nitrogen (e.g. denoted by H) might be labelled "Edinburgh H". While another axis might be labelled "Edinburgh L", where L denotes a Lower Level of Nitrogen. Because $I = 2$ the two Edinburgh axes coincide, as do the axes for all other sites.

Figure 6 displays such a triplot for the interaction array $\hat{\underline{\mathbf{Z}}}$. All IJ combinations of nitrogen-rate and trial-site are displayed by calibrated axes but only $J = 7$, rather than $IJ = 14$, distinct axes are necessary. We use the convention that the label "Edinburgh" denotes not only the site but also the high rate of nitrogen. The marker for the low rate of nitrogen in Edinburgh could be placed at the other end of the axis but it is superfluous. The markers on the axis are positive in the section between the label (e.g. Edinburgh) and the origin and negative away from the origin; the opposite holds for the implicit Edinburgh×low marker. All $K = 12$ varieties are displayed as points.

By projecting variety k onto the combined rate-site axes, triadic rank-two interactions can be read directly off the calibrations to give the estimation of the term for variety k and all combinations of levels of i and j. A 'projection circle' on the diameter determined by the point displaying the variety and through the origin, gives a convenient way of accessing all projections of the variety onto the $J = 7$ rate×trial axes together with their associated calibrations. Such projection circles have been introduced in the context of biplots in Gower and Hand (1996) and Gower et al. (2011), and in the context of triplots in Albers and Gower (2014).

Figure 6 shows this visualisation for the Blackman data where the point 'Cap' represents the variety Cappelle. Sites Begbroke, Trumpinton, and Earith give positive interactions, Boxworth about zero and sites Craftshill and Fowlmore give negative interactions between Cappelle and high levels of Nitrogen. The signs are reversed for interaction with low levels of Nitrogen. The intervals of calibration may be refined at will but here we give only a marker 10 grams per square meter. It is important to keep in mind when interpreting these triadic interactions that these are the values after main and biadditive effects (accounting for 98.07% of variation, see Table 2) have been partialed out: The triplot focuses on the remaining 1.93% of variation and large differences in the triplot denote, in this example, only relatively small differences on an overall level.

Note that (a) although this two-dimensional visualisation may look like a biplot it involves three factors and thus is really a triplot and (b) it remains valid when $I > 2$, though without the simplifactions of coincident axes, which might introduce visual overlad. Both Albers and Gower (2014) and Williams and Gardner-Lubbe (2016) provide examples of such a triplot with $I = 3$.

5 Discussion

Essentially, our approach is to adopt the usual linear models for representing main effects, two factor interactions and three factor interactions. The two factor interactions may be approximated by multiplicative bilinear terms and the three factor interactions may be approximated by multiplicative trilinear terms. In the bilinear case the approximations have standard least-square estimates, based on singular value decompositions, but in the trilinear case, we propose that the estimates be conditioned on the residuals from the saturated bilinear model. In principal, it would be possible to do a full unconditional least-squares solution but the conditional approach is easier and avoids difficulties with constraints. In the bilinear case identification constraints are not substantive but in the full trilinear case there is a troubling substantive interaction between the bilinear and trilinear parameter constraints. This problem is avoided when using the conditional method of analysis. The suggestion of applying a triadditive model to three-way residuals has also been made by van Eeuwijk and Kroonenberg (1998), who used a Tucker3 model rather than the Candecomp-Parafac model. Williams and Gardner-Lubbe (2016) use an orthogonal Parafac decomposition as basis for their visualisations and arrive at figures similar to Figure 3(a) on basis of geometric arguments.

We do not claim that the biadditive and triadditive models are substantive models per se, although in certain applications they could be. We make use of biadditive and triadditive models as a useful framework to base our visualisations on. A special virtue of biadditive models is the way that they lend themselves to simple biplots for visualising the interactions between rows and columns of the two classifying factors. This is particularly useful when biadditive interactions are adequately approximated in two dimensions and in this paper we have proposed how these biplots may be enhanced. It would be helpful if similar visualisations were available for triadditive interactions and, following Albers and Gower (2014), we demonstrate how two-dimensional triplots for rank-two tridimensional interaction tables may be formed, in which all three-dimensional tetrahedronal information is retained. When one factor is at two levels, some striking simplifictions occur, as is demonstrated in Section 4). When I, J, $K > 2$, there is a

risk of visual overload. Such overload can be reduced through smart choices, constructing parallel coordinate plots (such as Figure 5) for triplots and through interactivity. For instance, markers for calibrated axes could be displayed only when a certain axis is selected, and one could use tick boxes to select which of the IJ axes and K points should be shown. Finding out which approaches work best against visual overload is an interesting path for future research. Furthermore, additional smart choices w.r.t. calibration, (arbitrary) rotation and use of colour can enhance the interpretability (Blasius et al., 2009).

Rank two triplot displays in two dimensions seem to be at the bounds of practical utility. Attempts to visualise rank-three displays in three dimensions are not promising. Fortunately, as with biadditive biplots, it is the rank-two displays that are the most useful and rank two tridimensional visualisations show similar promise.

At the outset of this paper we drew attention to the adage of McCullagh and Nelder about interactions being predicated on their main effects and lower orders of interaction. Our approach of conditioning three-order interactions on main effects and two-factor interactions is in accord with the adage. Nevertheless, at several points in our discussion we have seen that main effects and lower order interactions may be ignored when fitting a higher-order interaction. Sometimes, but not always, it seems that, as with Tukey's model of non-additivity, additive terms may be absorbed in equivalent multiplicative parameterisations of the model. It seems to us that it is always wise to keep the McCullagh and Nelder adage in mind but there are occasions, especially with multiplicative relationships, when it is less persuasive.

Software

All computations have been performed in R, using self-written code (available upon request from the first author). For the Candecomp-Parafac decompositions the R-package ThreeWay (Giordani et al., 2014) has been used. For the correspondence analyses, the R-package ca (Nenadic and Greenacre, 2007) has been used.

Acknowledgements

Dr. Steffen Unkel (University of Göttingen, Germany) gave helpful comments on Section 2. The attendants of the TRICAP 2015 conference, Prof. Pieter Kroonenberg (Leiden University, The Netherlands) and Dr. Sugnet Lubbe (University of Cape Town, South Africa) in particilar, provided valuable feedback.

Appendices

A Calibrated biplots for biadditive interaction arrays

In the notation of Section 2.1 it is useful, especially when $R = 2$, to plot the rows of $\mathbf{c}_r$ $(r = 1, \ldots, I)$ to give I row-points and the rows of $\tilde{\mathbf{c}}_s$ $(s = 1, \ldots, J)$ to give J column-points. In this biplot, the inner-product determined by a pair of points, one from each set, gives a visualisation of the corresponding interaction. Here $\mathbf{c}_r$ and $\tilde{\mathbf{c}}_s$ derive from the SVD of $\underline{\mathbf{X}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}'$ and we set $\mathbf{c}_r = \mathbf{u}_r\boldsymbol{\Sigma}^{\alpha}$ and $\tilde{\mathbf{c}}_s = \mathbf{v}_s\boldsymbol{\Sigma}^{\beta}$ where usually $\alpha + \beta = 1$.

If we project $\tilde{\mathbf{c}}_s$ onto the vector $\mathbf{c}_r$ we find $\left[\boldsymbol{\Sigma}^{\beta}\mathbf{v}_s'\left(\mathbf{v}_s\boldsymbol{\Sigma}^{2\beta}\mathbf{v}_s'\right)^{-1}\mathbf{v}_s\boldsymbol{\Sigma}^{\beta}\right]\boldsymbol{\Sigma}^{\alpha}\mathbf{u}_r'$ which, when $\alpha + \beta = 1$ simplifies to

$$\left[\boldsymbol{\Sigma}^{\beta}\mathbf{v}_s'\left(\mathbf{v}_s\boldsymbol{\Sigma}^{2\beta}\mathbf{v}_s'\right)^{-1}\right]\mathbf{v}_s\boldsymbol{\Sigma}\mathbf{u}_r' = \left[\boldsymbol{\Sigma}^{\beta}\mathbf{v}_s'\left(\mathbf{v}_s\boldsymbol{\Sigma}^{2\beta}\mathbf{v}_s'\right)^{-1}\right]z_{rs}. \tag{A1}$$

In (A1), only the interaction z_{rs} depends on r so all points $r = 1, \ldots, I$ are collinear on an axis with direction given by the term of (A1) given in square brackets. It follows that $\left[\boldsymbol{\Sigma}^{\beta}\mathbf{v}_s'\left(\mathbf{v}_s\boldsymbol{\Sigma}^{2\beta}\mathbf{v}_s'\right)^{-1}\right]$ may be used to calibrate the axis with values $\mu_1, \mu_2, \mu_3, \ldots$ usually chosen with an even calibration interval κ as $\mu, \mu \pm \kappa, \mu \pm 2\kappa, \ldots$. Setting $\mu = 0$ gives the scale for $z_{\mu s}$. If we set $\mu = a_r$ the markers include the main effect of the ith main effect a_r, so giving the combined effects of the main effect and interactions of r with all the columns s. Note that this merely requires a cosmetic change to the markers and not any extra calculation.

Similarly, all rows $r = 1, 2, \ldots, I$ may be shown as calibrated axes and if we project $\mathbf{c}_r$ onto the vector $\tilde{\mathbf{c}}_s$ all columns $s = 1, 2, \ldots, J$ may be shown as axes calibrated in terms of $\left[\boldsymbol{\Sigma}^{\alpha}\mathbf{u}_r'\left(\mathbf{u}_r\boldsymbol{\Sigma}^{2\alpha}\mathbf{u}_r'\right)^{-1}\right]$.

Note that the marker for z_{rs} occurs twice, once on $\mathbf{c}_r$ and once on $\tilde{\mathbf{c}}_s$. Furthermore, the distances of the two markers from the origin are unequal. It would be elegant to arrange equal scaling but we have not succeeded and believe it to be impossible.

B Automatic ordering of the parallel axes

In constructing parallel coordinate plots as Figures 4 and 5, the ordering of the axes usually is irrelevant (unless the corresponding factor is at some ordinal level). In that case, visual information might be gained by rearranging the axes optimally.

In total, $J!$ orderings are possible and, by excluding mirrorings ('ABCD' yields the same information as 'DCBA'), there are $J!/2$ orderings to choose between.

This appendix explains an automated procedure to do so, based on correspondence

analysis (CA). CA is similar to principal component analysis, but for nominal-labelled data.

Let $\mathbf{Y}$ be the $J \times K$ table with the projections for the K varieties on the J sites (either with or without main effects). The goal is to rearrange the J columns optimally; i.e. such that projections on adjacent axes are as close as possible. Since correspondence analysis is designed for non-negative data, we shift $\mathbf{Y}$ such that all values are non-negative, i.e. through $\mathbf{Y}' = \mathbf{Y} - \min \mathbf{Y}$. Since the row sums of $\mathbf{Y}$ are zero (since the average interaction per site is zero), and hence the row sums of $\mathbf{Y}'$ are equal, some simplifications with respect to general correspondence analysis are possible, although the gain in computation speed is negligible for small values of J (such as in Section 4).

The simplified algorithm is as follows:

1. Compute $\mathbf{M} = \mathbf{S} - \mathbf{w}_J \mathbf{w}'_K$, where $\mathbf{S} = \mathbf{Y}' / \sum\sum y'_{jk}$, $\mathbf{w}_J$ is the $J \times 1$ vector of row weights with equal entries $1/J$, and $\mathbf{w}_K$ is the $K \times 1$ vector with entries $\sum_j y'_{jk} / \sum_{jk} y'_{jk}$;
2. Perform a SVD on $\mathbf{M}$: $\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}'$ under the restrictions $\mathbf{U}'\mathbf{U} = J\mathbf{I}$ and $\mathbf{V}'\mathrm{diag}(\mathbf{w}_K)\mathbf{V} = \mathbf{I}$;
3. Compute $\mathbf{F}_J = \mathbf{U}\mathbf{\Sigma}$;
4. Rearrange the J rows of $\mathbf{Y}$ according to the ordering in the first column of $\mathbf{F}_J$.

In Step 4, one could rearrange the rows ascending or descending, which yields two visualisations that are one another's mirror image.

References

C. J. Albers and J. C. Gower. A contribution to the visualisation of three-way arrays. *Journal of Multivariate Analysis*, 132:1–8, 2014.

C. J. Albers, J. C. Gower, and H. A. L. Kiers. Rank properties for centred three-way arrays. In F. Mola, C. Conversano, and M. Vichi, editors, *Classification, (Big) Data Analysis and Statistical Learning*, Studies in Classification, Data Ananlysis, and Knowledge Organization. New York: Springer, 2017.

L. Araújo. Seleçao e análise dos modelos PARAFAC e Tucker e gráfico triplot com applicaçáo em interaçáo tripla. doctoral thesis. university of Sao Paulo, Brazil, 2009.

J. A. Blackman, J. Bingham, and J. L. Davidson. Response of semi-dwarf and conventional winter wheat varieties to the application of nitrogen fertilizer. *Journal of Agricultural Science*, 90:543–550, 1978.

J. Blasius, P. H. C. Eilers, and J. C. Gower. Better biplots. *Computational Statistics and Data Analysis*, 53:3145–3158, 2009.

R. Bro and M. Jakobsen. Exploring complex interactions in designed data using GE-MANOVA. color changes in fresh beef during storage. *Journal of Chemometrics*, 16 (6):294–304, 2002. doi: 10.1002/cem.722.

J. D. Carroll and P. Arabie. Multidimensional scaling. *Annual Review of Psychology*, 31:607–649, 1980.

J. D. Carroll and J. J. Chang. Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart-Young' decomposition. *Psychometrika*, 35:283 – 319, 1970.

C. Coombs. *A theory of data*. New York: John Wiley, 1964.

L. C. A. Corsten and A. C. Van Eijnsbergen. Multiplicative effects in two-way analysis of variance. *Statistica Neerlandica*, 26:61–68, 1972.

J. B. Denis and J. C. Gower. Biadditive models. *Biometrics*, 50:310 – 311, 1994.

M. d'Ocagne. *Coordonnées parallèles et axiales: Méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles*. Paris: Gauthier-Villars, 1885.

H. D. Gauch. *Statistical analysis of regional yield trials: AMMI Analysis of Factorial Designs*. Amsterdam: Elsevier, 1992.

P. Giordani, H. A. L. Kiers, and M. A. del Ferraro. Three-way component analysis using the R package ThreeWay. *Journal of Statistical Software*, 57, 2014.

H. F. Gollob. A statistical model that combines features of factor analytic and analysis of variance techniques. *Psychometrika*, 33:73–115, 1968.

J. C. Gower. The analysis of three-way grids. In P. Slater, editor, *Dimensions of Intra Personal Space, vol. 2, The Measurement of Intra Personal Space by Grid Technique*, pages 163–173. Chicester: Wiley, 1977.

J. C. Gower and D. J. Hand. *Biplots*. London: Chapman and Hall, 1996.

J. C. Gower, P. J. F. Groenen, and M. van de Velden. Area biplots. *Journal of Computational and Graphical Statistics*, 19:46 – 61, 2010.

J. C. Gower, S. Lubbe, and N. Le Roux. *Understanding biplots*. Chicester: Wiley, 2011.

M. Greenacre. *Correspondence analysis in practice*. London: Chapman & Hall, second edition, 2007.

R. A. Harshman. Foundations of the PARAFAC procedure: Models and methods for an 'explanatory' multi-mode factor analysis. *UCLA Working Papers in Phonetics*, 16:1 84, 1970.

A. I. Inselberg. *Parallel Coordinates: Visual Multidimensional Geometry and its Applications*. Springer, 2009.

H. A. L. Kiers. Towards a standardized notation and terminology in multiway analysis. *Journal of Chemometrics*, 14:105–122, 2000.

P. M. Kroonenberg. *Applied Multiway Data Analysis*. Hoboken, New Jersey: Wiley, 2008.

P. M. Kroonenberg and T. H. A. van der Voort. Multiplicatieve decompositie van interacties bij oordelen over werkelijkheidswaarde van televisiefilms. *Kwantitatieve methoden*, 8:117–144, 1987.

P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Boca Raton, Florida: Chapman & Hall/CRC, 2nd edition, 1989.

J. J. Meulman, A.J. van der Kooij, and W. J. Heiser. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data. In D. Kaplan, editor, *The SAGE handbook of quantitative methodology for the social sciences*. London: Sage, 2004.

O. Nenadic and M. Greenacre. Correspondence analysis in R, with two- and three-dimensional graphics: the ca package. *Journal of Statistical Software*, 20, 2007.

C. R. Rao. *Advanced Statistical Methods in Biometric Research*. New York: John Wiley & Sons, 1952.

E. Schmidt. Zur theorie der linearen und nichtlinrearen Integralgleichungen. 1. Teil: Entwicklung willkürlicher Funktionen nach System vorgeschriebener. *Mathematische Annalen*, 63:433 – 476, 1907.

A. K. Smilde, R. Bro, and P. Geladi. *Multi-way analysis with applications in the chemical sciences*. Hoboken, New Jersey: John Wiley & Sons, 2004.

J. M. F. ten Berge. Simplicity and typical rank results for three-way arrays. *Psychometrika*, 76:3 – 12, 2011.

M. E. Timmerman and H. A. L. Kiers. Three-mode principal component analysis: choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology*, 53:1–16, 2000.

L. R Tucker. Some mathematical notes in three-mode factor analysis. *Psychometrika*, 31:279–311, 1966.

J. W. Tukey. One degree of freedom for non-additivity. *Biometrics*, 5:232–242, 1949.

F. A. van Eeuwijk and P. M. Kroonenberg. Multiplicative models for interaction in three-way ANOVA, with applications to plant breeding. *Biometrics*, 54:1315 – 1333, 1998.

D. Williams and S. Gardner-Lubbe. Visualising three-way arrays. *Chemometrics and Intelligent Laboratory Systems*, 158:180–186, 2016.

Tables

Trial site	abbreviation
Craftshill	Cra
Begbroke	Beg
Fowlmere	Fow
Trumpington	Tru
Boxworth	Box
Earith	Eea
Edinburgh	Edi

Variety	abbreviation
Cappelle	Cap
Ranger	Ran
Huntsman	Hun
Templar	Tem
Kinsman	Kin
Fundin	Fun
Durin	Dur
Hobbit	Hob
Sportsman	Spo
TJB259.95	259
TJB325.464	325
Hustler	Hus

Table 1: Overview of the trial sites (left) and varieties of wheat (right) of the Blackman data set, as well as the abbreviations used in later visualisations.

Factor	SS	df	% of total
A (rate of nitrogen application)	125078	1	4.84
B (trial site)	1854207	6	71.72
C (variety of wheat)	196211	11	7.59
AB	221481	6	8.57
AC	8021	11	0.31
BC	130411	66	5.04
$r = 1$	60961	(16)	
$r = 2$	42642	(14)	
$r = 3$	12623	(12)	
$r = 4$	8334	(10)	
$r = 5$	3799	(8)	
$r = 6$	2053	(6)	
ABC	49812	66	1.93
Total	2585224	167	

Table 2: ANOVA-breakdown of Blackman's data. The SS for the rows with specific values for r are obtained via (7). The corresponding degrees of freedom are obtained via the rule of thumb explained in Section 2.2. (Note that, since $df_A = 1$, no similar breakdown for the AB and AC interaction is possible.)

Rank S	Fit (%)	Increment
1	35.40	35.40
2	63.10	27.70
3	78.62	15.52
4	88.89	10.27
5	97.74	8.85
6	100.00	2.26

Table 3: Candecomp-Parafac approximations to the three-way interaction ABC for different ranks S for Blackman's data.

Figures

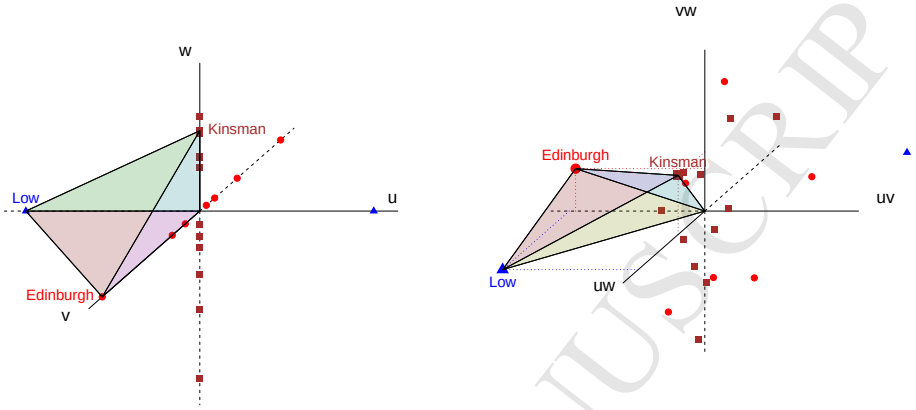


Figure 1: Rank $R = 1$ (left) and $R = 2$ (right) fits to the triadditive terms for Blackman's data. Blue triangles refer to Factor A (the levels of nitrogen), red circles to Factor B (trial sites) and brown squares to Factor C (varieties). For the $R = 1$ fit, all points lie on orthogonal axes, for the $R = 2$ fit, they all lie on orthogonal planes. The tetrahedra corresponds to the interaction "low nitrogen × Edinburgh × Kinsman".

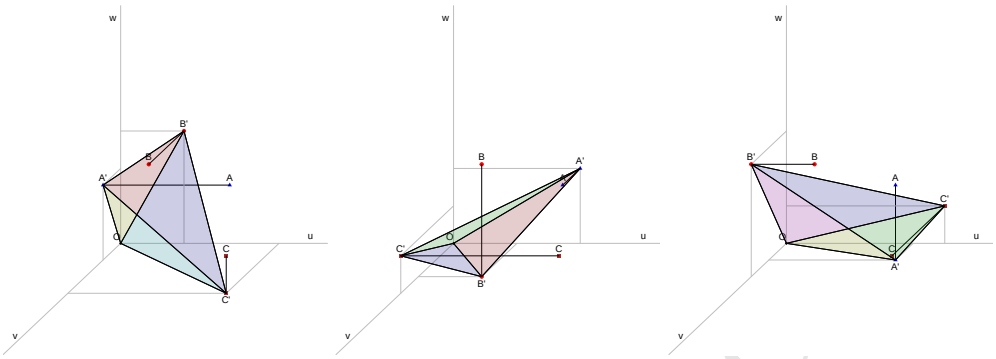


Figure 2: A demonstration of a three-way interaction for the rank $R = 3$ fit to triadditive terms, for a constructed example with conveniently chosen coordinates. All levels of all factors now have coordinates that are not restricted to (orthogonal) axes nor planes. The three points A, B, C, are projected onto the vw, uw and uv planes, respectively. Subsequently, the polygon OA'B'C' is constructed (left). Similarly, polygons are constructed for projections onto uw, uv and vw (middle) and uv, vw and uw (right). The interaction ABC is proportional to the sum of the volumes of the three tetrahedra thus obtained.

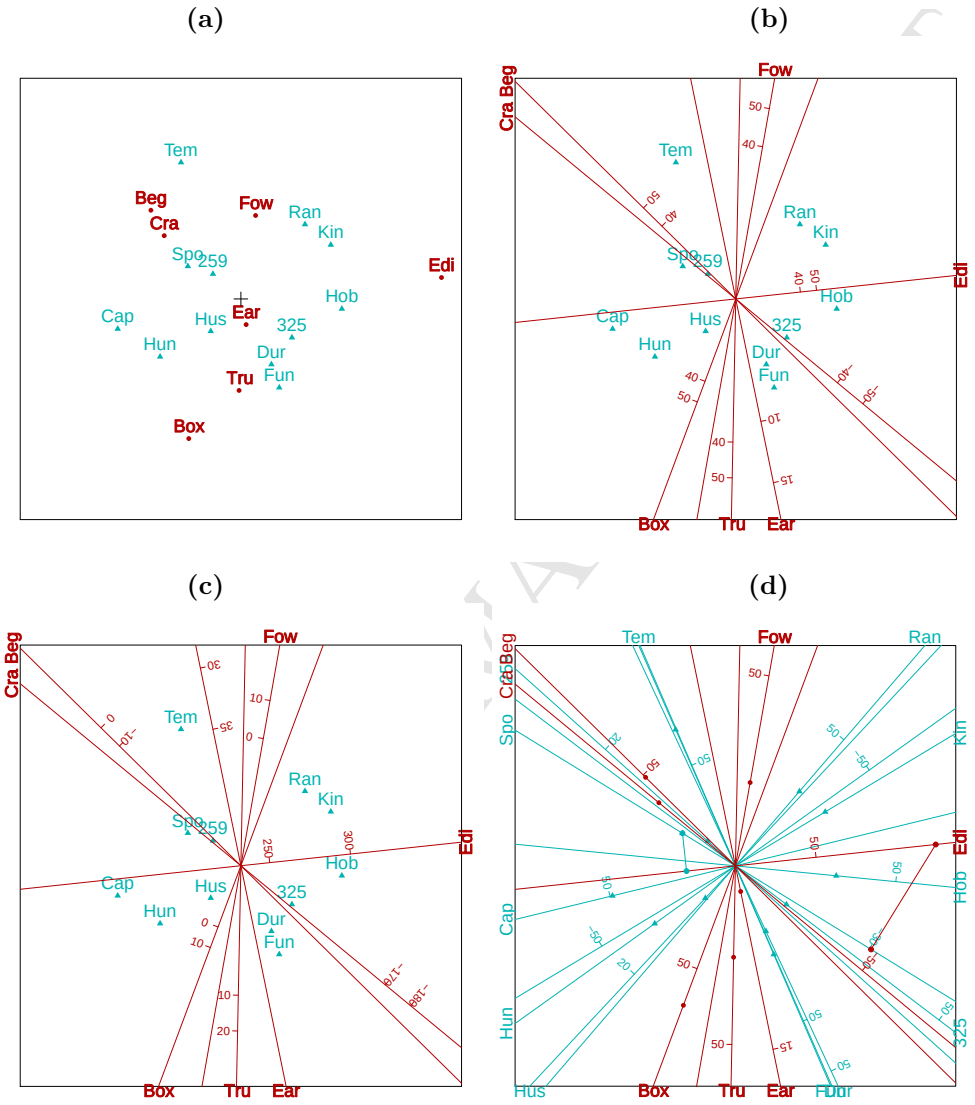


Figure 3: Visualisation of the rank $R = 2$ approximation to the biadditive interaction between factors B and C. First (a) a regular biplot is given (with + indicating the origin; trial locations are denoted by '·' and varieties by a triangle), followed by a biplot where trial sites have been replaced by calibrated axes; where calibration is done with $\mu = 0$ (b) and $\mu = b_j$ (c). Finally, panel (d) shows a biplot where both varieties and trial sites are represented by calibrated axes. Abbreviations in bold font correspond to trial sites. See Table 1 for the full labels for the abbreviations.

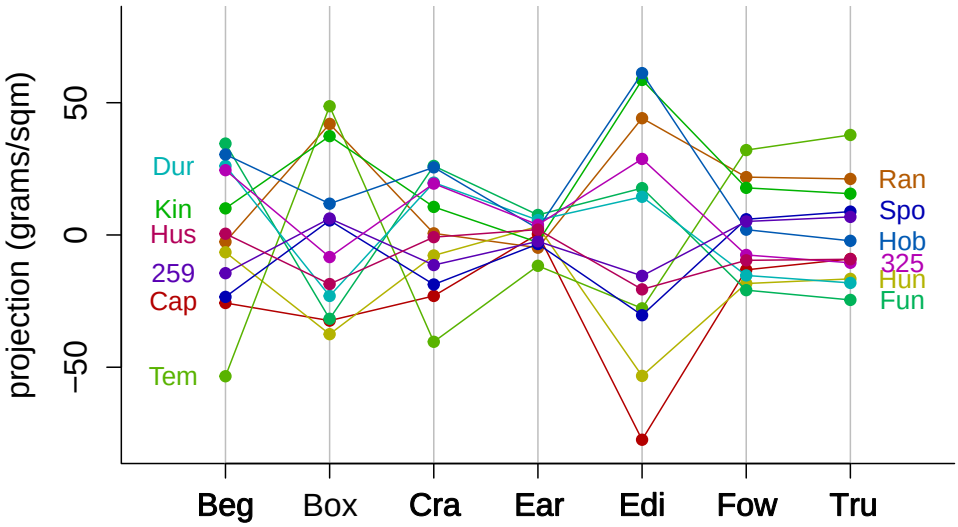


Figure 4: For all 7 trial sites the projections of the varieties (with $\mu = 0$) are given in this single-axis diagram. A single calibrated axis applies to all sites. Abbreviations in bold font correspond to trial sites. See Table 1 for the full labels for the abbreviations.

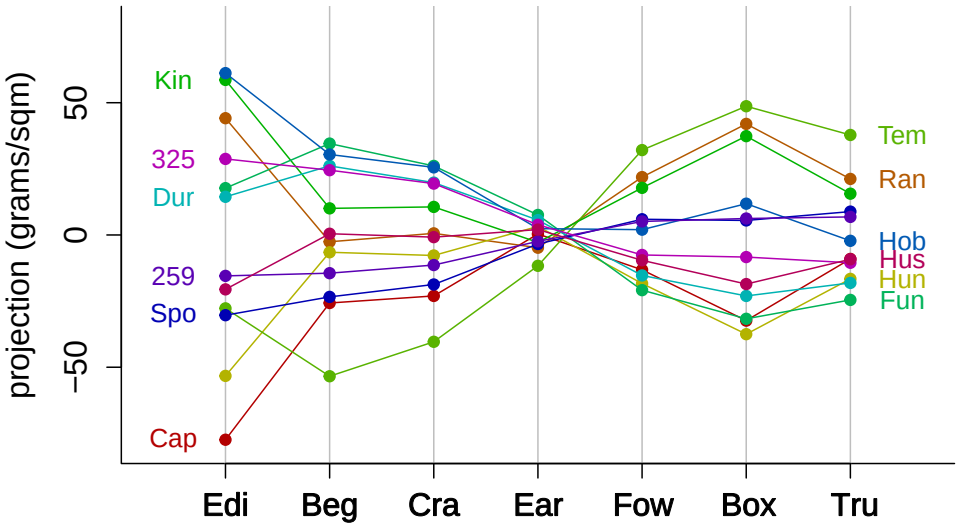


Figure 5: A similar visualisation as Figure 4, now with the ordening of sites according to the correspondence analysis algorithm outlined in Appendix B.

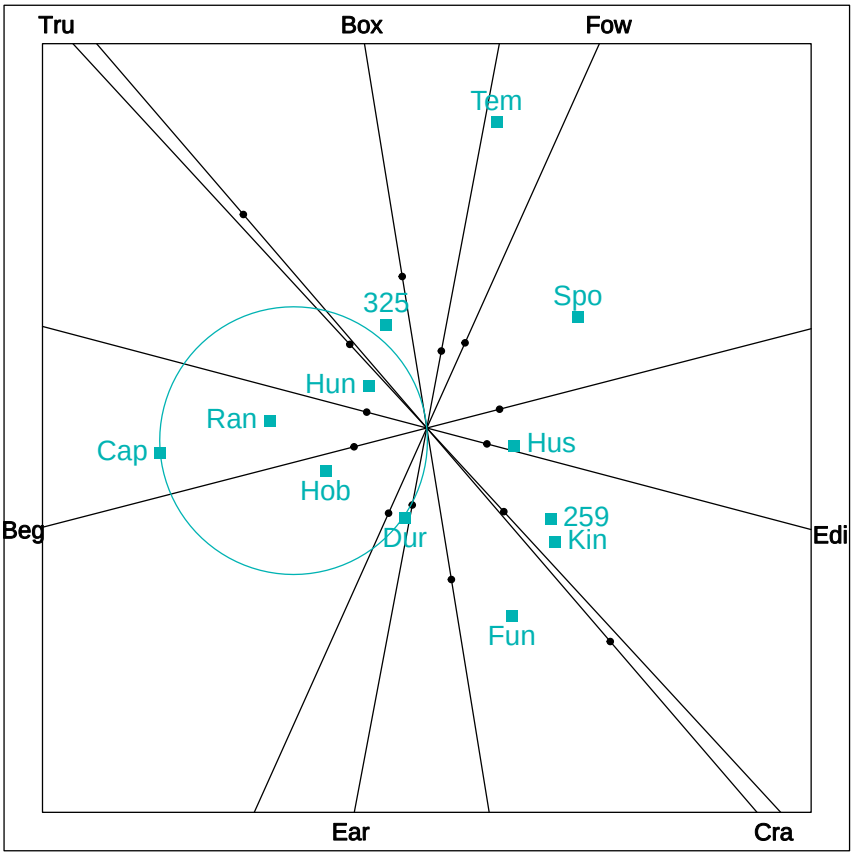


Figure 6: Triplot. Axes represent rates and sites. Since $I = 2$, the axes for low and high rates coincide. Site labels are placed at the positive end of the 'high'-axis. The signs are reversed for predicting interactions to the low rate of nitrogen. A single positive and negative marker is shown on each axes; these correspond to 10 grams per square meter grain yield. A projection circle through Cappelle cuts the axes at the calibrations corresponding to the seven calibration points giving the rank-two triadic interactions. These are positive or negative, depending on whether they occur on the same or opposite side of the origin as the site label. Abbreviations in bold font correspond to trial sites. See Table 1 for the full labels for the abbreviations.